GPR-Based Surrogate Model for Probabilistic Fatigue Safety Factor Prediction in Stepped Shafts

Sai Patil
COEP Technological University, Pune

Abstract

This study develops a Gaussian Process Regression (GPR) surrogate model to predict the fatigue safety factor of stepped shafts, addressing the computational expense and meshing constraints of traditional Finite Element Analysis. A dataset of 314 high-fidelity ANSYS simulations was generated using Latin Hypercube and custom sampling across five parameters: minor diameter (d), major diameter (D), fillet radius (r), bending moment (M), and loading ratio (R). Trained with a Matern $\nu = 2.5$ kernel, the model achieved exceptional accuracy (R2 = 0.9905, RMSE = 0.187) and robust generalization (5-fold CV, R2 = 0.971 ± 0.022). Permutation importance revealed a physically consistent parameter ranking (d > M > R > D > r) while predicted trends for R aligned with Goodman mean-stress correction theory. Beyond point predictions, GPR provides quantified uncertainty estimates, enabling risk-aware design decisions and targeted simulation refinement. This work demonstrates that Bayesian surrogate modeling can effectively replace repetitive FEA fatigue analyses, offering a computationally efficient, uncertainty-aware methodology that accelerates design exploration and enhances reliability assessment under constrained simulation environments.

Contents

1 Introduction

Stepped shafts are critical mechanical components prone to fatigue failure at geometric discontinuities, making accurate safety factor (SF) prediction essential for reliable design. Traditional Finite Element Analysis (FEA), while highly accurate, is computationally intensive and often restricted by software licensing or mesh limits, hindering efficient design exploration. To overcome these bottlenecks, this study develops a Gaussian Process Regression (GPR) surrogate model that rapidly predicts fatigue SF while providing quantified prediction uncertainty, making it particularly effective for data-scarce engineering applications. Using a curated dataset of 314 ANSYS-simulated design points generated via a combination of custom sampling and Latin Hypercube Sampling, the model captures the complex, nonlinear interactions among five key parameters: minor diameter (d), major diameter (D), fillet radius (r), bending moment (M), and loading ratio (R). By demonstrating how Bayesian machine learning can effectively replace repetitive high-fidelity fatigue analyses, this work offers a computationally efficient, uncertainty-aware methodology that accelerates mechanical design optimization and enhances reliability assessment.[3]

2 Problem Definition

The geometry of the shaft is defined using the following parameter: d – minor/smaller shaft diameter, D – major/larger shaft diameter, r – fillet radius[1] at the joint section. One end is fixed (zero displacement) and a bending moment M is applied at the free end of the minor shaft section. Another important parameter involved in defining the outcome is R, the loading ratio R which is defined as the ratio of minimum stress to maximum stress (R = $\sigma_m in/\sigma_m ax$). R characterizes the nature of cyclic loading — fully reversed at R = -1, pulsating at R = 0, and mean-stress-dominated at R = 0.3 — and directly influences fatigue damage accumulation rate. The output we have considered is the safety factor (SF) [2] that we get from ANSYS Mechanical fatigue analysis tool. The input ranges considered in this problem are given as follows :$r \in [1, 5]$, $d \in [20, 50]$, $D \in [25, 70]$, $M \in [106, 947]$ Nm, $R \in \{-2, -1, -0.5, 0, 0.3\}$. The geometry choices and combinations are significantly limited by the meshing node limit (of 128k) applied by ANSYS on the free student version. The only constraint used in this geometry model is that the smaller shaft diameter should be smaller than the larger shaft diameter or simply put, D>d.

3 Data Generation

The data used to train the model is generated using different tools offered by ANSYS Mechanical software. SpaceClaim to design the geometric model, static structural tool to apply constraint and loading on the geometry, fatigue analysis tool to extract safety factor SF and fatigue life as outputs. Design points were generated using a combination of custom sampling and Latin Hypercube Sampling (LHS), using ANSYS Workbench's inbuilt DOE tool. A total of 314 datapoints are used to train the model. The datapoints are distributed across different R values, R = -1 is overrepresented relative to other values due to additional sampling conducted to improve coverage in the fully reversed loading regime. All simulations were performed using Structural Steel as the shaft material, with fatigue properties defined by the ANSYS default S-N curve for Structural Steel. The obtained data was cleaned using checks D>d and the calculated safety factor should be greater than 0, SF>0. One point with $D < d$ (physically invalid geometry) was removed, resulting in a final dataset of 314 points.

Table 1: Distribution of simulation points across loading ratio R

R value	-2	-1	-0.5	0	0.3
Datapoints	50	115	50	50	50

4 GPR Model

GPR is a non-parametric, Bayesian machine learning technique used for regression and probabilistic inference. Unlike traditional models that learn a fixed set of parameters, GPR treats the regression function as a distribution over functions, allowing it to model complex, nonlinear relationships without assuming a specific functional form.[3] GPR outputs a predictive mean along with an uncertainty estimate, quantifying the model's confidence at each prediction point. For engineering design, a probabilistic range of safety factor is more actionable than a single fixed value, it allows designers to account for model confidence when making decisions near failure thresholds. GPR is preferred here over neural networks and polynomial fitting for two reasons. Neural networks require large datasets to generalize well, which is impractical given the computational limitations of ANSYS simulations. Polynomial fits cannot capture the nonlinear interactions between geometric and loading parameters across a five-dimensional input space. The Matern ν=2.5 kernel was selected as it models twice-differentiable functions, appropriate for smooth physical systems such as stress and fatigue responses. Kernel parameters were optimized automatically by maximizing the log marginal likelihood[3] during training. The model was implemented using GaussianProcessRegressor from scikit-learn.[4]

Input features were scaled using StandardScaler prior to fitting, as GPR is sensitive to the relative scale of input variables. The target variable was normalized by setting `normalize_y=True` to handle the wide range of SF values in the dataset. Data was split into an 80/20 train/test ratio with `random_state=42` for reproducibility.

5 Results

5.1 Predictive Accuracy

The parity plot (equivalence plot) is a plot of predicted values against actual/real values. Here, it is the GPR predicted values of SF on y-axis and actual ANSYS simulated/calculated values of SF on x-axis. The ideal case where all the predicted values are the same as the actual values (GPR predicted SF = ANSYS simulated SF) is depicted by a dashed diagonal line passing from the origin. The GPR predicted values are depicted as points (with a variance of $\pm 2\sigma$ on the y-axis) on the plot. The points closer to the diagonal have less error and the farther the points get from the diagonal, higher the error. R2 and RMSE are model evaluation metrics used to understand a model. R2 is a measure of how well a model understands the variation in the dataset. We got R2 = 0.9905 showcasing that the GPR model is excellent at understanding the distribution of data across the dataset. In other words, the model is able to explain 99% of the variation in the dataset. RMSE or Root Mean Squared Error, is a standard metric used to evaluate the accuracy of regression models by measuring the average magnitude of the errors between predicted and actual values. We got RMSE = 0.187, meaning that the average error the model makes while predicting SF is 0.187, which is considerably less considering the range of SF. Another way to verify the results of a model is to use cross validation. Here, we have used 5-fold cross validation, meaning across 5 different training and testing splits, the model was evaluated on R2 metric. The result we got is, R2 = 0.971 ± 0.022, meaning across 5 different train-test splits the model is able to understand and explain 97% of the variation with a standard deviation of 0.022 across folds.

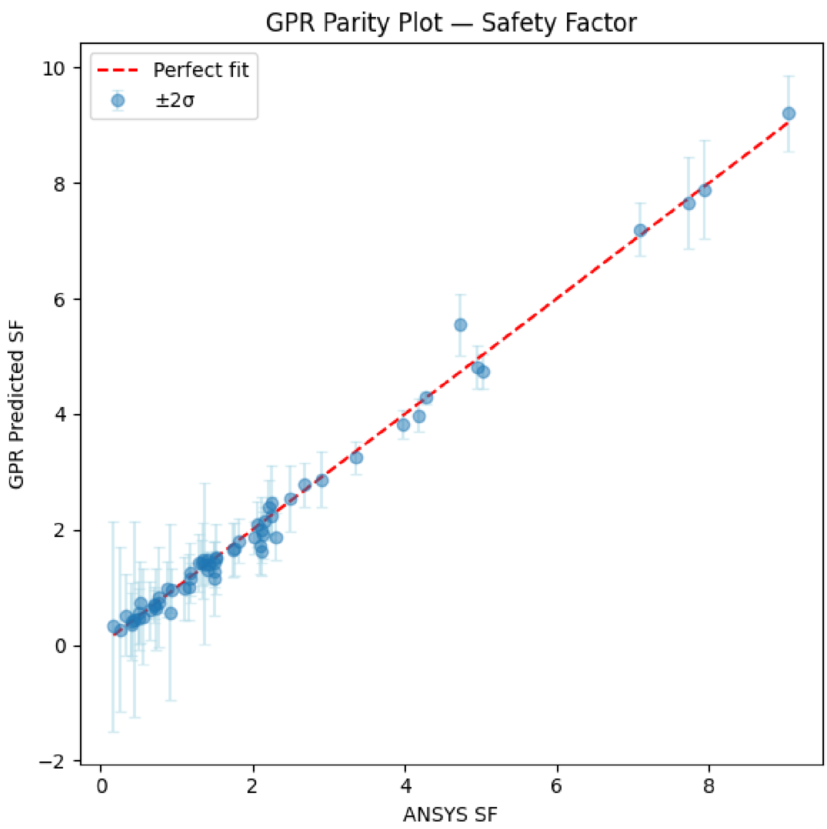

Figure 1: GPR parity plot — predicted vs ANSYS SF with $\pm 2\sigma$ uncertainty bounds

5.2 Feature Importance

For a better understanding of the effect of each parameter on the output safety factor SF, a permutation importance chart is helpful. We get the following ranking sequence : $d > Moment > R > D > r$ The parameter d (smaller shaft diameter) is the most important parameter which affects the fatigue life of the shaft the highest. As stress is inversely proportional to safety factor, from the formula relation between stress and other parameters given as, $\sigma = \frac{32M}{\pi d^3}$. From this relation we directly know that, stress is directly proportional to moment M and stress is inversely proportional to the cube of d, hence small changes in d result in considerable variations in stress and indirectly safety factor SF . The relation also explains why moment M is second in the ranking sequence. The loading ratio R ranks third, confirming that it has meaningful independent influence on SF beyond geometry and applied load alone. This justifies its inclusion as a fifth input parameter. D and r are ranked lower, though r's physical influence through stress concentration factor Kt is likely underestimated due to its narrower variation range in the dataset.

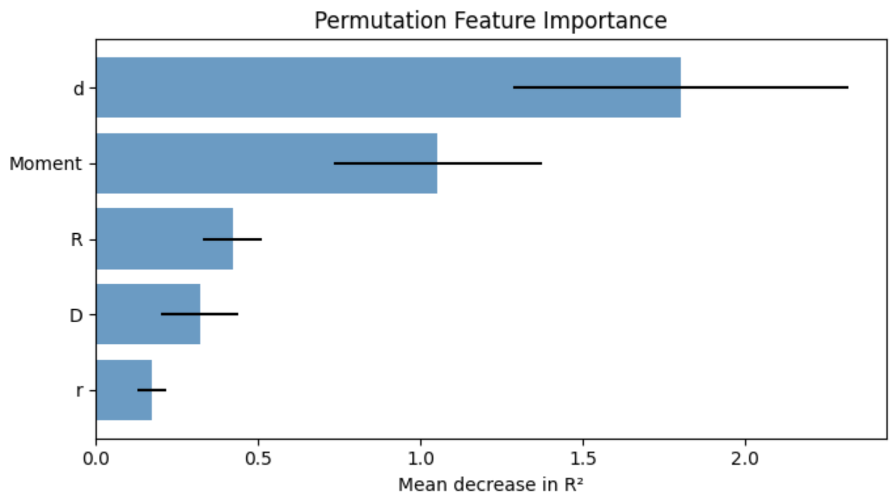

Figure 2: Permutation feature importance — mean decrease in R^2

5.3 Effect of Loading Ratio R

To better understand the effects of loading ratio R on the safety factor SF, a plot of predicted SF vs R was plotted. The plot shows the variation of predicted SF against loading ratio R under different cases such as small shaft high load, medium shaft medium load, large shaft low load. As R increases, SF shows different variations for different cases. For small shaft high load case the SF is almost always below SF=1 meaning it is most likely to fail under fatigue loading. For medium shaft medium load case the SF varies close to SF=1, it increases as R increases from -2 to 0.3 . For large shaft low load case, the SF stays always above the SF=1 line. All the cases show a common characteristic of increasing SF as R increases. The shaded bands around the different case lines is the model uncertainty. It is mainly caused by less training data for the specific case and less neighbouring datapoints. Fatigue failure is driven by stress amplitude, a parameter defined as the amount of swing a load goes through in a single loading cycle. Experiments show that if you have the same stress amplitude, a tensile mean stress makes fatigue worse and a compressive mean stress makes it better. The observed increase in predicted SF with increasing R is consistent with Goodman mean stress correction theory, where reduction in stress amplitude at higher R values outweighs the penalty from increased mean stress.[2]

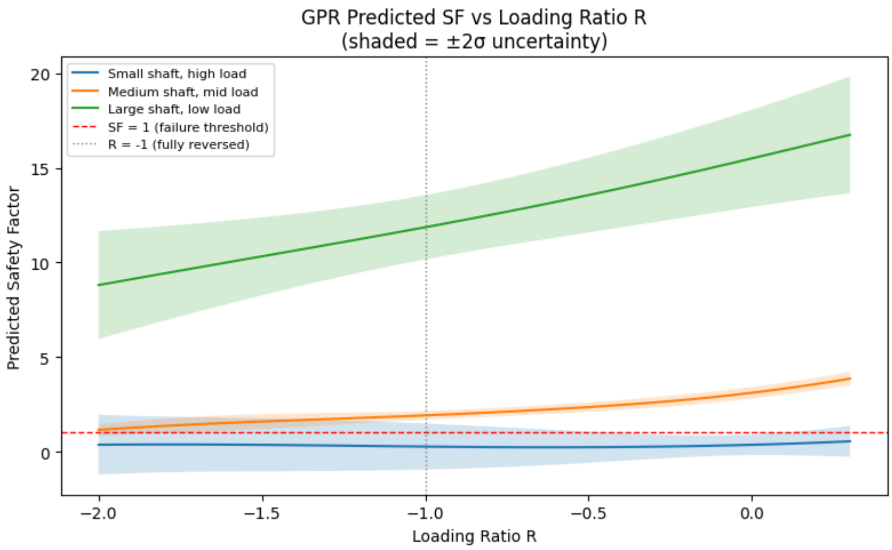

Figure 3: GPR predicted SF vs loading ratio R for three shaft cases with $\pm 2\sigma$ uncertainty bands

6 Discussion

The uncertainty estimates provided by GPR serve as more than just prediction intervals, they also signify the need for more training data in the specific region. With more training data the model understands the relation better in specific regions hence improving the predictions and decreasing the uncertainty to a comfortable, workable range. The surrogate model is trained on a single material (Structural Steel) and does not generalize to other materials without retraining on material-specific fatigue data. The GPR model predicts a negative SF for multiple edge cases, which is a limitation and requires further refinement. The problem can also be addressed with a positivity constraint, which is not supported by scikit-learn's GaussianProcessRegressor. A log-transform of the target variable SF could enforce positivity and improve model behaviour at extreme input combinations; this approach was not implemented here as the problem was only limited to the minor edge cases where the data is sparse and for the majority of the input space the model behaved correctly. The sparse training data in larger size of shafts was caused due to the meshing node limit (128K) enforced by ANSYS on the student version of ANSYS Mechanical. The problem is to be addressed by using a paid version of ANSYS Mechanical with higher meshing node limit.

7 Conclusion

This study presents a GPR-based model for predicting the safety factor SF of a stepped shaft with 5 input parameters. For a training dataset of 314 points the model was able to get $R2 = 0.9905, RMSE = 0.187$ and a 5-fold cross validation $R2 = 0.971 \pm 0.022$. The permutation feature importance chart showed a physically consistent ranking sequence, $d > Moment > R > D > r$. R the loading ratio ($R = \sigma_min/\sigma_max$) is the third parameter in the 5 parameters input space which affected SF the most. The surrogate model reduces the need for repeated ANSYS simulations during design exploration, while GPR's uncertainty estimates provide actionable confidence intervals rather than single-point predictions.

References

[1] Pilkey, W. D., & Pilkey, D. F. (2008). *Peterson's Stress Concentration Factors* (3rd ed.). Wiley.

[2] Budynas, R. G., & Nisbett, J. K. (2020). *Shigley's Mechanical Engineering Design* (11th ed.). McGraw-Hill.

[3] Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

[4] Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.